Object localization in the presence of a strong heterogeneous background in fluorescent tomography

Pouyan Mohajerani,* Ali A. Eftekhar, and Ali Adibi

Department of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Drive NW, Atlanta, Georgia 30332, USA
**Corresponding author: pouyan@ece.gatech.edu*

Received October 19, 2007; revised March 11, 2008; accepted April 1, 2008;
posted April 8, 2008 (Doc. ID 88835); published May 30, 2008

We propose a method for object localization in fluorescent tomography (FT) in the presence of a highly heterogeneous background. Existing approaches typically assume a homogeneous background distribution; thus, they are incapable of accurately accounting for the more general case of an unconstrained, possibly heterogeneous, background. The proposed method iteratively solves the inverse problem over a solution space partitioned into a background subspace and an object subspace to simultaneously estimate the background and localize the target fluorescent objects. Simulation results of this algorithm applied to continuous-wave FT demonstrate effective localization of target objects in the presence of highly heterogeneous background distributions. © 2008 Optical Society of America

OCIS codes: 170.3010, 170.0170, 170.3880, 170.7050.

1. INTRODUCTION

The development of activated and targeted fluorescent molecular probes [1] has facilitated *in vivo* study and monitoring of certain molecular processes several millimeters under the skin. Self-quenching fluorescent probes, for example, have been designed to specifically target cancer-associated proteases [2,3]. These probes are activated upon encountering the targeted enzyme, at which time the interaction cleaves a peptide bond within the probe, thereby activating the attached fluorophores. Coated quantum dots are another technology under investigation for *in vivo* imaging. Available with a wide variety of coatings, these particles are covered with particular conjugate binding sites that will specifically interact only with the target molecules [4]. The high binding affinity of fluorescent probes has made them an attractive tool for cancer research and for the development and monitoring of drugs such as protease inhibitors [5,6].

With the opportunities afforded by the probe technologies, the ability to image and quantify the concentration of fluorophores deep within tissue (>1 mm) is of great importance. The study and understanding of the propagation of light in highly scattering biological tissues, especially in the near-infrared (NIR) wavelengths (700–900 nm) where tissue has the lowest absorption, has paved the way for the development of fluorescent tomography (FT) techniques for *in vivo* three-dimensional imaging of fluorescent molecules using optical surface measurements [7–11].

While FT has the potential to achieve submillimeter resolution [8] and picomole sensitivity [7], several factors affect the accurate localization and quantification of fluorescent objects. One especially important obstacle is the presence of strong background fluorescence in tissue. There are two primary sources that contribute to the background fluorescence: Autofluorescence of the tissue and fluorescence from excess probes (i.e., the extrinsic probes not uptaken by the targeted regions) [12]. The background signal can overwhelm the desired fluorescent signal from targets in relatively deep locations, thus adversely affecting object localization and quantification [13,14].

Several methods have been proposed to handle the background fluorescence. One intuitive method is the differential approach where an estimate of the background is obtained by acquiring measurements before and after the probe administration. The applicability of this approach to *in vivo* scenarios is limited due to the considerable time span between the two measurements during which the physiology within the body may change considerably. Also, it is not possible to estimate the background introduced after the administration of probes using this method. More practical methods either preprocess the measurements to estimate and subtract the background signal or account for the background fluorescence during the reconstruction [12,14].

Most methods proposed for reduction of background fluorescence assume a homogeneous distribution for the background. However, the homogeneous model is not always adequate for modeling the heterogeneous background common to *in vivo* scenarios; therefore, more versatile techniques that do not rely on a homogeneous model are of interest [13].

In this work, we propose a method to improve the performance of FT in localization and quantification of fluorescent inclusions in the presence of strong and highly heterogeneous background fluorescence. We use the difference between the spatial characteristics of the distri-

butions of object and background fluorophores to define two sets of basis functions for representing them. The background distribution, though highly heterogeneous, usually follows smooth spatial variations, while the object fluorophores are sparsely distributed [15–17]. The validity of such object–background modeling of the fluorophore distribution is discussed later in Section 6. We employ a set of spatially smooth basis functions for modeling the background such that an arbitrary background distribution is approximated as a nonnegative linear combination of these basis functions. The set of all distributions spanned by this basis set is referred to as the background subspace. The object distribution is modeled by an object subspace consisting of sparse distributions. The object and background distributions are estimated simultaneously by iteratively solving the inverse problem over the object and background subspaces. At each iteration, the solution over each subspace is used to improve the estimation over the other. The results of the simulation of a FT system with a continuous-wave laser source for a cylindrical phantom are presented to verify the ability of the proposed method to effectively separate the object and background signals for cases in which the homogeneous model does not provide sufficient representation of the background signal and, therefore, fails in object reconstruction.

The remainder of this paper is divided into six sections: Section 2 describes the linear system model of FT, Section 3 defines the object and background subspaces, Section 4 presents our method for separation of the object and background distributions using these subspaces, and Section 5 presents detailed simulation results for one typical case. Section 6 discusses the performance of the proposed method and the validity of the assumptions using further simulation results, and Section 7 is the conclusion.

Note on Notations. Throughout this paper, matrices are denoted by bold, capital letters. The n-dimensional space of real numbers (positive real numbers) is denoted by $\Re^n(\Re^{+n})$. The L_0, L_1, and L_2 norms of $x \in \Re^n$ are denoted as $\|x\|_0$, $\|x\|_1$, and $\|x\|$, respectively. For $x,y \in \Re^n$, the inner product of x and y is denoted by $\langle x,y\rangle$. The zero vector in $\Re^n$ is shown as $\mathbf{0}_n$. The symbols $\equiv$ and $\approx$ are interpreted as "defined as" and "approximated by," respectively. The symbol $\triangle$ is used to mark the end of a proof. For a given finite set S, the number of the elements of S is shown by $|S|$. For a given vector X, diag(X) represents a diagonal matrix with X at the main diagonal, and X^T represents the transpose of the vector X. The subsets of $\Re^n$ are denoted using bold, capital letters with an upper bar, as in $\bar{\mathbf{X}}$. Finally, the term subspace, as used in this paper, refers to a subset of $\Re^n$ and does not necessarily imply a vector subspace in the algebraic sense.

2. LINEAR SYSTEM MODEL OF FLUORESCENT TOMOGRAPHY

In this section, we review the mathematical model describing the relationship between surface optical measurements and the distribution of fluorophores. The propagation of the NIR light in a highly scattering tissue can be modeled by the diffusion approximation (DA) [18]. We use finite-element modeling (FEM) for discretization of DA for numerical modeling [19,20] using a tetrahedral mesh with K nodes.

Let X be a $K\times 1$ vector representing the fluorophore concentration (defined as the product of the absorption cross section and the quantum yield of fluorophores) at the K mesh nodes. Also let M be the $N_dN_s\times 1$ vector of measurements for the N_s source and N_d detector locations. As previously described in [15], the relationship between the measurement vector M and the fluorophore distribution vector X can be expressed by the following linear model:

$$M = \mathbf{Z}X + \eta, \tag{1}$$

where the $N_dN_s\times K$ matrix $\mathbf{Z}$ is the system matrix (derived as a function of the tissue and measurement geometries and the optical properties of the tissue, using FEM modeling of DA) and η is an $N_sN_d\times 1$ vector representing the additive measurement noise, which is modeled by a Gaussian random process (see [21] for noise modeling). We assume the optical properties and, thus, the system matrix $\mathbf{Z}$ are known prior to reconstruction of the fluorophore distribution. While diffuse optical tomography techniques can be used to estimate these parameters [21], a common approach consists of assuming a uniform optical absorption and scattering equal to their typical values in tissue [22]. Proper normalization of the emission measurements by the excitation measurements is then applied to mitigate the effect of the optical heterogeneities [13,22,23].

3. MODELING THE DISTRIBUTION OF THE OBJECT AND BACKGROUND FLUOROPHORES

The distributions of the fluorophores in the background and in the targeted areas have different spatial properties: Background fluorophores, even if heterogeneous, have a rather smooth distribution, while the fluorophores in the targeted areas are distributed in a set of small isolated objects that are attributed to neoplastic tissues [17,24]. The differences in their respective spatial distributions allow their representation by two different sets of basis functions: One for smooth distributions and one for sparse, isolated distributions. (Note: The terms "basis function" and "basis distribution vectors" are used here interchangeably.) It should be noted that while such modeling of the fluorophore distribution as composed of the two object and background components does include a wide range of applications, it does not include all practical cases. The limitations and validity of this modeling approach are further discussed in Section 4. Throughout the development of the proposed method, however, we assume this model for the fluorophore distribution.

As discussed in Section 2, the distribution of fluorophores in tissue is modeled by a $K\times 1$ vector of nonnegative real values, whose kth element denotes the concentration of fluorophores on the kth mesh node. Let $\bar{X}$ be the space of all possible fluorophore distributions, which in the most general form is $\Re^{+K}$. The fluorophore distribu-

tion vector X may be written as a combination of the object distribution X_o and the background distribution X_b as

$$X = X_b + X_o. \tag{2}$$

Next we present a framework for representation of the object and background distributions.

A. Representation of the Background Distribution

The smooth spatial variations of the background fluorescence allows for its representation by a set of smooth basis functions. Appendix A presents a criterion for the spatial smoothness of a given distribution. Let P denote the number of the basis functions used in the background distribution expansion. The background distribution, denoted by the K-dimensional vector X_b, can generally be represented by a nonnegative P-dimensional vector, Y_b, in which each element of Y_b corresponds to a nonnegative expansion coefficient for one of the basis functions. The background distribution is approximated as a weighted summation of the basis functions as

$$X_b \approx \tilde{X}_b = \sum_{i=1}^{P} Y_b(i) B_i, \tag{3}$$

where B_i is a K-dimensional vector representing the value of the ith basis function on the K mesh nodes and $Y_b(i)$ denotes the ith element of the vector Y_b.

A set of smooth basis functions is constructed using the sinc function, which is advantageous due to its maximal smoothness and minimal high-frequency content. Consider the following nonnegative function consisting of the main lobe of the one-dimensional sinc function:

$$g^{\omega}:\Re \to \Re^{+}; \quad g^{\omega}(x) = \begin{cases} \dfrac{\omega \sin(x/\omega)}{x} & |x| \leq \pi\omega \\ 0 & \text{else} \end{cases}, \tag{4}$$

where the real number ω controls the level of the smoothness of this function. The P basis functions, β_i^{ω}, are then defined as

$$B_i^{\omega} = [g^{\omega}(\|q_1 - c_i\|), \ldots, g^{\omega}(\|q_K - c_i\|)] \quad (i = 1, \ldots, P), \tag{5}$$

where q_j is a three-dimensional vector denoting the spatial position of the jth mesh node. The three-dimensional vectors $c_i (i = 1, \ldots, K)$ denote the center positions of the basis functions and are chosen to span the volume of the tissue. Finally, we define the background subspace as

$$\bar{X}_b = \left\{ \sum_{i=1}^{P} \alpha_i B_i^{\omega} | \alpha_i \geq 0 \right\}. \tag{6}$$

It is shown in Appendix A that the smoothness of the elements of $\bar{\mathbf{X}}_b$ is controlled by the choice of ω and P. Sufficient smoothness can be achieved due to the high smoothness of the function g^{ω}.

B. Representation of the Object Distribution

In contrast to the background, the object distribution consists of a few isolated areas within the tissue. For a given measurement geometry and noise level, FT has a limited resolution in the detection of fluorescent objects; therefore, a fluorescent object up to a certain size can be sufficiently modeled by a fluorescent point source, regardless of the actual shape of the object [25]. A combination of point sources can be used to represent larger objects. With this in mind, the object distribution can be modeled as a sparse vector, namely, one with few nonzero elements (a method for object localization using this property is proposed in [15]). Accordingly, the object subspace is defined as the set of all sparse distributions. One method to ensure a sparse solution is to limit the maximum number of nonzero elements. In this sense, the object space, denoted by $\bar{\mathbf{X}}_o$, is given as

$$\bar{\mathbf{X}}_o = \left\{ \sum_{i \in I} a_i \delta_i | a_i \in \Re^{+}, |I| \leq N_o, I \subset \{1, \ldots, K\} \right\}, \tag{7}$$

where the delta function δ_i is equal to 1 only on the ith mesh node and 0 elsewhere. The members of $\bar{\mathbf{X}}_o$ have, at most, N_o nonzero elements. This approach, while promoting the convergence of the proposed iterative method presented in Section 4, limits the maximum number of nonzero elements; therefore, placing an upper limit on the number and sizes of objects that can be represented in the object subspace $\bar{\mathbf{X}}_o$.

Sparseness can also be imposed while solving the linear problem by minimizing the L_1 norm of the solution vector—a method referred to as basis pursuit (BP) [26]. This approach does not limit the maximum number of nonzero elements and, further, yields a convex optimization problem. A combination of these two approaches for representation of sparse distributions is used in the iterative method proposed in the next section.

4. PROPOSED METHOD FOR JOINT OBJECT AND BACKGROUND ESTIMATION

As explained in the previous section, the object and background distributions can be each approximated by a linear combination of a set of basis functions. We estimate the expansion coefficients for the object and background distributions through constrained minimization of the distance between the predicted and the actual measurements. Since different constraints should be applied to solve for the expansion coefficients of the object and background distributions based on the corresponding basis function sets, joint optimization of the object and background distribution parameters is a nonlinear problem.

We propose a method to iteratively estimate the object and background distributions by solving the inverse problem over the object and background subspaces. The result of the estimation over each subspace is used to improve the result of the estimation over the other. The iterations are repeated until the overall estimation error, defined as the distance between the predicted and the actual measurements, converges to within a preset bound.

The sequence of the steps comprising the proposed method is presented in Table 1. Let $X_o^{(i)}$ and $X_b^{(i)}$ denote the estimated object and background distributions, respectively, at the ith iteration. The initial estimations are set to zero in step (a). In step (b) we find the best candidate in the background subspace for the current estimation of the background signal, i.e., $M - M_o^{(i)}$. The resulting optimization problem for finding the expansion coeffi-

Table 1. Proposed Method for Estimation of Object and Background Distributions

(a) Initialization; $i=0$, $X_b^{(0)}=\mathbf{0}_K$, and $X_o^{(0)}=\mathbf{0}_K$. Start in the OMP mode.

(b) Estimate the background distribution given the current estimation of the object distribution;

$$X_b^{(i+1)}=\arg\min_{x\in\bar{\mathbf{X}}_b}\|\mathbf{Z}x-(M-M_o^{(i)})\|=\sum_{j=1}^{P}a_jB_j^{\omega},$$

where

$$(a_1,\ldots,a_P)=\arg\min_{a_1,\ldots,a_P\geq 0}\left\|\mathbf{Z}\sum_{j=1}^{P}a_jB_j^{\omega}-(M-M_o^{(i)})\right\|.^{a}$$

(c) Estimate the object distribution given the current estimation of the background distribution;

If in the OMP mode, then

$$X_o^{(i+1)}=\arg\min_{x\in\bar{\mathbf{X}}_o,\|x\|_0\leq N_o}\|\mathbf{Z}x-(M-M_b^{(i+1)})\|=\sum_{j\in I}b_j\delta_j,$$

where

$$((b_j)_{j\in I},I)=\arg\min_{b_j\geq 0,|I|\leq N_o}\left\|\mathbf{Z}\sum_{j\in I}b_j\delta_j-(M-M_b^{(i+1)})\right\|,$$

else

$$X_o^{(i+1)}=\arg\min_{x\in\Re^{+K}}\|\mathbf{Z}x-(M-M_b^{(i+1)})\|+\lambda\|x\|_1.^{b}$$

(d) Calculate the estimation error; $E_{i+1}=\|M-M_b^{(i+1)}-M_o^{(i+1)}\|$.

(e) If the estimation E_i is not stable,[c] then $i=i+1$ and go to (b).

(f) If in the OMP mode, then switch to the BP mode and go to (b), else finished.

[a] The estimated background and object signals at iteration i are given as $M_b^{(i)}=\mathbf{Z}X_b^{(i)}$ and $M_o^{(i)}=\mathbf{Z}X_o^{(i)}$, respectively.

[b] The value of the constant λ is adjusted based on the system matrix and the noise power. For a discussion, see [26].

[c] Stability of the estimation error is assessed based on the amount of variations in the last three values of E_i.

cients a_i is a quadratic problem. Next in step (c) we find the best candidate in the object subspace based on the current estimation of the object signal $M-M_b^{(i+1)}$.

Optimization over the object subspace implies searching for the sparsest eligible solution, namely, one with a minimum number of nonzero elements. As explained in Subsection 3.B, the sparseness of the object distribution can be enforced in two ways: By explicitly applying an upper limit to the number of the nonzero elements in the object distribution or by minimizing the L_1 norm during the minimization of the estimation error. For convergence reasons explained in Appendix C, we first use the former approach and, after convergence, switch to the latter approach for enforcing sparseness.

First, using the former approach, we try to estimate the object distribution as a vector with, at most, N_o nonzero elements in step (c) of Table 1. While for a given set of indices of nonzero elements, denoted by I, finding the optimal coefficients, denoted by b_i, is straightforward, optimal selection of I necessitates an exhaustive search, which is computationally prohibitive. To avoid this complexity, we use a suboptimal search approach based on orthogonal matching pursuit (OMP) [27] to suboptimally estimate the object distribution as a sparse vector (the details of OMP are explained in Appendix B for comprehensiveness and clarity). This mode of operation is identified as the "OMP mode" in Table 1.

Then, after convergence in the OMP mode, the algorithm switches to the method of BP based on the L_1 minimization, referred to as the "BP mode" in Table 1, to relax the restriction imposed on the maximum number of the nonzero elements during the OMP mode. The optimization term in the BP mode is a weighted combination of the L_1 norm of the solution vector and the fidelity term (i.e., the distance between the predicted and the actual measurements). The algorithm terminates after convergence in the BP mode.

After each iteration, the distance between the actual measurement vector and the predicted measurement vector, given as $E_i=\|M-M_b^{(i)}-M_o^{(i)}\|$, is evaluated in step (d). In step (e), the stability of this error is evaluated and used as the convergence criterion. Various stability tests can be used for this purpose. As an example, we evaluate the variations in the last few E_i's and determine the convergence accordingly. A discussion regarding the convergence of the iterative algorithm presented in Table 1 for the two modes is presented in Appendix C.

5. SIMULATION RESULTS

To demonstrate the efficiency of the proposed method, we present here in detail the simulation results for a case involving a highly heterogeneous background distribution. A cylindrical phantom with a diameter of 4 cm, a height of 6 cm, and an optical absorption and scattering of, respectively, $0.2\,\mathrm{cm}^{-1}$ and $10\,\mathrm{cm}^{-1}$ (typical for bulk tissue) is used for all simulations. A total of 18 sources and 18 detectors are placed on the surface of the cylinder as shown in Fig. 1 (similar measurement geometries have been experimentally used in other studies [28]). The extrapolated zero boundary condition [29] is used to model the air–tissue boundary, assumed at the entire phantom surface, and a mesh of 3403 nodes and 14,702 tetrahedral ele-

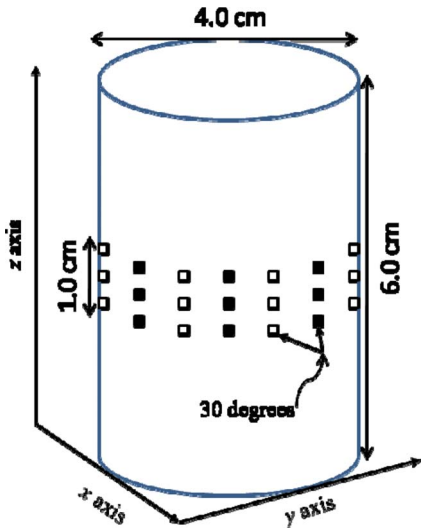

Fig. 1. (Color online) Spatial configuration of the simulation case study discussed in Section 5. The cylinder represents the phantom, and the filled (open) square marks indicate the source (detector) locations.

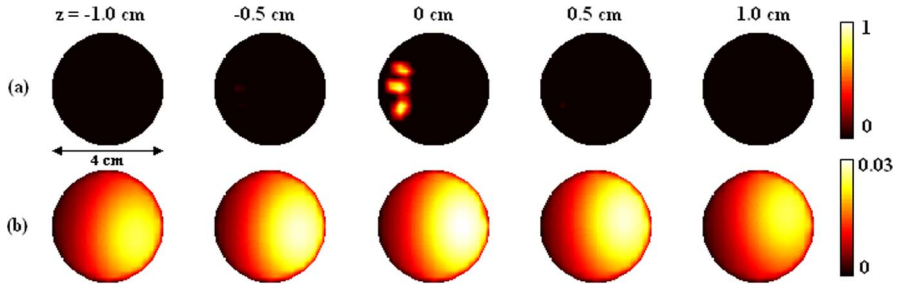

Fig. 2. (Color online) The z-slice images of (a) the object and (b) the background fluorophore distribution for the simulation case study discussed in Section 5. The fluorophore distribution inside the phantom consists of three spherical objects, each with a diameter of 5 mm and a background distribution consisting of two Gaussian functions, each with a standard deviation of 15 mm centered at $z=-5$ mm and +5 mm.

ments is used to discretize the tissue volume. This simulated FT setup is assumed to operate in the continuous-wave mode.

The simulation results are presented for the case of a fluorophore distribution consisting of two spherical objects, each with a diameter of 5 mm (of interest in early cancer detection [24]), and a background distribution consisting of two Gaussian distributions with a standard deviation of 15 mm (the Gaussian distribution has been used for modeling the distribution of nonspecific fluorescent agents [13]). The objects are located on the $z=0$ plane about 1 cm under the phantom surface, and each spans over three mesh nodes. The object distribution is modeled as having a concentration equal to 1 over a total of nine nodes and 0 elsewhere. The z-slice plots of the object and background distributions for $z=-1.0$, -0.5, 0, 0.5, and 1.0 cm are depicted in Figs. 2(a) and 2(b), respectively.

The relative object-to-background concentration ratio is adjusted such that the object and background signals have equal powers. The resultant tumor-to-background ratio (TBR, defined as the concentration of fluorophores on the object divided by the average concentration in the background) is approximately 125. Note that even for such a high TBR, because of the larger quantity of the fluorophores in the background and the deep location of the objects, the background signal still significantly affects the measurements. A measurement noise with a signal-to-noise ratio (SNR) of 20 dB, modeled by a Gaussian random process, is also added to the measurements.

In the OMP mode, we consider at most one nonzero node for the object distribution (i.e., $N_o=1$). Using the definition presented in Eqs. (4)–(6), the background subspace is represented by a set of modified sinc basis functions with $\omega=1/3\ \mathrm{cm}^{-1}$, centered on 220 locations inside the phantom volume (i.e., $P=220$).

The z-slice plots of the estimated object and background distributions are depicted in Figs. 3(a) and 3(b). As can be seen from Fig. 3, the proposed method provides a fairly accurate estimation of both the object and background distributions. For comparison, we have also estimated the background assuming a homogeneous model by restricting the background subspace to the uniform distributions, namely, $\{\alpha\tau|\alpha\geq 0\}$, where the distribution τ is equal to 1 over all mesh nodes except the extrapolated boundary nodes. As can be seen from the results presented in Fig. 3(c), estimation using the homogeneous model generates a highly erroneous solution. Finally, in

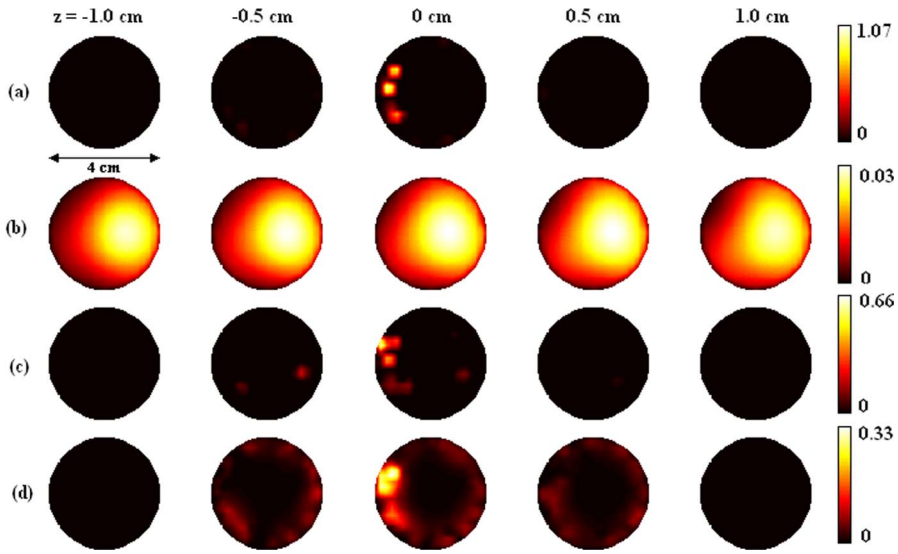

Fig. 3. (Color online) The z-slice images of (a) the object and (b) the background distributions estimated using the proposed approach after 25 iterations; (c) the object distribution as estimated using an approach based on the assumption of homogeneity of the background distribution; (d) the object distribution estimated using Tikhonov regularization without any background compensation.

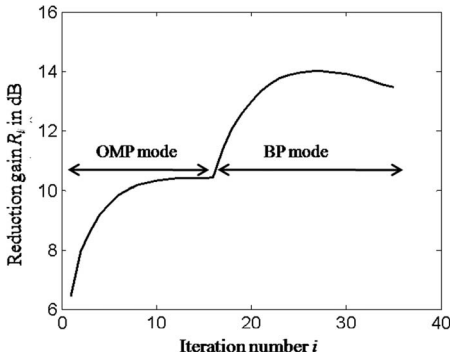

Fig. 4. Background reduction gain in decibels versus the iteration number for the simulated case. The algorithm switches from the OMP mode to the BP mode at the 16th iteration. The high background reduction achieved in the OMP mode is maintained in the BP mode, where the object estimation is further improved.

Fig. 3(d) we show the result of reconstruction using Tikhonov regularization [10] without any compensation for the background signal.

As a performance measure, we define the reduction gain at the ith iteration as

$$R_i = 20 \log_{10}\left(\frac{\|M_b\|}{\|M_b - M_b^{(i)}\|}\right), \tag{8}$$

where M_b and $M_b^{(i)}$ are the actual and estimated background signals, respectively, in the ith iteration. The reduction gain is used as a measure of the fidelity in the estimation of the background signal. For instance, a 20 dB reduction gain indicates a 90% power reduction of the background fluorescence signal in the measurements.

The reduction gain achieved using the homogenous model is about 3 dB, compared to the 14 dB reduction gain achieved using the proposed method. Figure 4 shows the reduction gain R_i calculated at each iteration of the proposed method. Sufficient modeling of the object distribution requires at least three nonzero components, that is, one for each object. Although the object subspace in the OMP mode is limited to vectors with, at most, one nonzero element, the algorithm achieves a relatively high reduction gain in this mode. In the BP mode, given the already relatively accurate background estimation, the L_1 minimization further improves the object estimation. Moreover, the reduction gain in the OMP mode tends to increase as the algorithm proceeds.

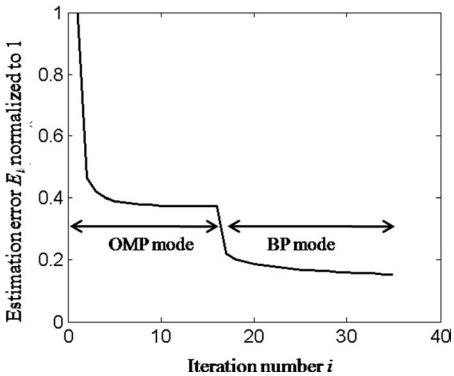

Fig. 5. Estimation error E_i versus the iteration number the simulated case of Fig. 1. The estimation error is nonincreasing in both the OMP mode and the BP mode.

Figure 5 depicts the estimation error E_i. The algorithm was repeated 15 times to converge in the OMP mode and then 20 times to converge in the BP mode. The estimation error follows a nonincreasing trend, which shows the convergence of the algorithm in both modes. This simulation scenario verifies the ability of OMP to achieve a fair initial estimation of the background and the helpfulness of the L_1 minimization to further improve the object estimation. The simulation case study described in Section 5 lasted about 100 s in the OMP mode and 500 s in the BP mode. Simulations were performed in MATLAB on a Pentium 4, 3.4 GHz processor with 1 Gbyte of RAM.

6. DISCUSSION

The method proposed in this paper aims at resolving the object and background signals by projection of the measurement vector into two relatively exclusive subspaces, namely, the object and background subspaces. While significant improvements over traditional methods for reducing the background fluorescence can be achieved, especially for highly heterogeneous distributions of background fluorescence as demonstrated in Section 5, several factors affect the performance of the proposed method. In this section we discuss such factors and analyze the performance of the method under various conditions using further simulation results. Specifically, we first review the assumptions made in the development of the proposed method and discuss their validity in practical scenarios. Next, we discuss the factors affecting the performance of the proposed method in separating the object and background signals. Finally, we investigate object and background estimation using the proposed method in the presence of modeling mismatch error due to optical perturbations.

For convenience and comprehensiveness, the following notations and definitions are used throughout this section. We define $M_o = \mathbf{Z}X_o$ and $M_b = \mathbf{Z}X_b$ as, respectively, the object and background signals (X_o,X_b, and $\mathbf{Z}$ are the object distribution, the background distribution, and the system matrix, respectively). Also, we define $\bar{\mathbf{M}}_o = \mathbf{Z}\bar{\mathbf{X}}_o$ and $\bar{\mathbf{M}}_b = \mathbf{Z}\bar{\mathbf{X}}_b$ (matrix multiplication performed on every vector in the corresponding subspace) as the object signal subspace and the background signal subspace, respectively [$\bar{\mathbf{X}}_o$ and $\bar{\mathbf{X}}_b$ denote the object and background subspaces defined in Eqs. (6) and (7), respectively]. Finally, let M_o^o, M_o^b, M_b^b, and M_b^o denote the projections of object and background signals (as specified by the corresponding subscripts) onto the object and background subspaces (as specified by the corresponding superscripts).

A. Object–Background Model for Fluorophore Distribution

The main assumption made in the development of the proposed method is that the fluorophores in tissue are either highly localized in target locations or are otherwise distributed diffusely as background fluorescence. In other words, the background fluorophores are assumed to be

distributed with a significantly lower level of spatial variations than the object fluorophores. This model for fluorophore distribution well represents the *in vivo* cases where the administered probes have a high binding affinity only to the targeted molecules, as is the case of various fluorescent probes engineered to target and/or activate upon encountering cancer-related enzymes. However, such a classification might not be possible when no specific objects can be distinguished in the fluorophore distribution or when the object and the background distributions have similar spatial characteristics and are, therefore, partly or completely indistinguishable using the object and background subspaces for a given set of subspace parameters (an example of such a case is later demonstrated in Fig. 9 and discussed in Subsection 6.B).

For the purpose of investigating the validity of this model, it is helpful to study the capability of the object and background subspaces to represent objects with different sizes, as illustrated in Fig. 6. In this figure, the solid (dashed) curves show the error of representing the object signal using the object (background) subspaces for different subspace parameters. This error is calculated for the object subspaces as follows. For a given object size and location, the projection of object signal M_o into the object subspace, i.e., M_o^o, is found and the error, in percentage, is calculated as $100\times\|M_o-M_o^o\|/\|M_o\|$. The depicted value in Fig. 6 is the average of this error taken over many object locations for each object size. Note that the projection into the object subspace is done using the OMP. Since OMP is suboptimal, the actual error is expected to be slightly smaller than the depicted value for object subspaces with two or more point sources (i.e., $N_o>1$). The methodology is the same for background subspaces.

As is observed in Fig. 6, the smaller objects are better represented with the object subspaces than the larger objects. Furthermore, more point sources generally result in a better representation. The background subspaces show poor representation for small objects, which implies a good separation between the object and the background signal subspaces. However, as objects become larger they better conform to the diffuse distribution model and are thus better represented by the background subspace. It should be also noted that the larger the background subspace parameter ω [refer to Eqs. (4)–(6)], the smoother the background subspace and the larger the object representation error. For significantly large objects there is strong cross talk between the object and the background subspaces; a point further illustrated in the next subsection.

As seen in Fig. 6, for the measurement geometry of Fig. 1, objects that are smaller than 1 cm in diameter (of interest in early cancer detection) are fairly well represented using the object subspaces, while they are poorly approximated by the background subspaces, which implies a low level of cross talk for such sizes. Note that the results presented in Fig. 6, as well as in other figures in this section, are obtained using the configuration of Fig. 1 and are expected to change according to the measurement geometry. For instance, increasing the number of source–detector pairs is expected to increase the distance between the signal subspaces, $\bar{\mathbf{M}}_o$ and $\bar{\mathbf{M}}_b$, thus reducing the cross talk between the object and the background signals and improving the performance of the proposed method in resolving the two distribution components.

As predicted by Fig. 6, large objects can be modeled as a collection of point sources; a point further illustrated by a simulation case presented in Fig. 7. Here the result of the reconstruction of a large object (2 cm in diameter) and a small object (0.5 cm in diameter) coexisting in a phantom, Fig. 7(a), in the presence of the background fluores-

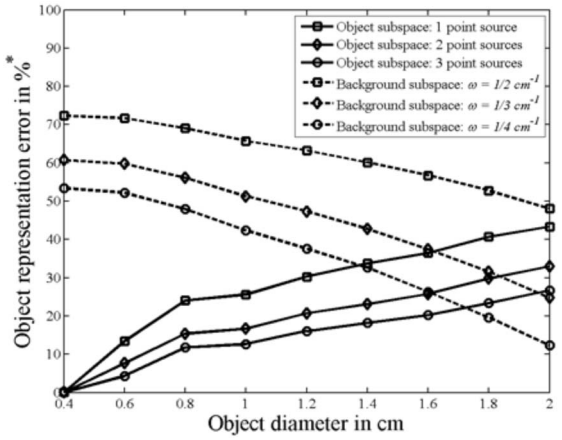

Fig. 6. Object representation error versus object size using object and background subspaces for the measurement geometry of Fig. 1. For a given object location, the error is given as $100\times\|M_o-M_o^p\|/\|M_o\|$, where M_o and M_o^p are the object signal and its projection, respectively, into the corresponding subspace. The depicted error is averaged over many object locations.

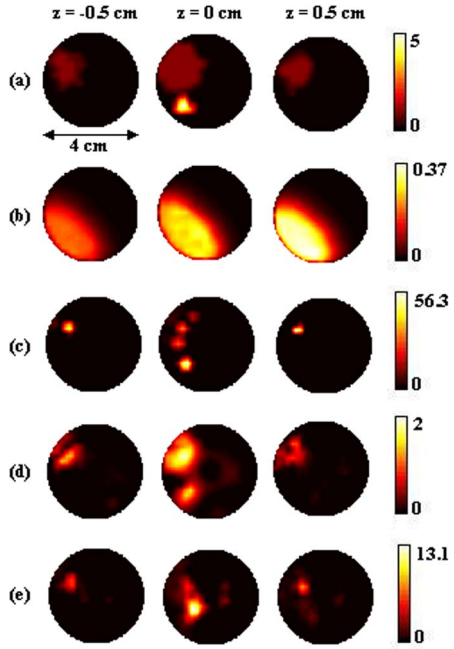

Fig. 7. (Color online) Reconstruction of large and small objects coexisting in the phantom. (a) Actual object distribution consisting of a large object (2 cm in diameter) and a small object (0.5 cm in diameter) and (b) the actual background distribution. As seen in (c), the large object is reconstructed as a few point sources using the proposed method. In (d) the object is reconstructed using L_2 minimization after subtracting the background estimation obtained using the proposed method. The object distribution estimated using the uniform background model and L_2 minimization is shown in (e) for comparison.

cence depicted in Fig. 7(b), is shown in Fig. 7(c). As seen, the larger object is estimated by a few point sources with a higher concentration. In Fig. 7(d) we have also shown the result of object reconstruction based on the minimization of the L_2 norm of the solution vector (i.e., the Tikhonov regularization, also referred to here as L_2 minimization) after subtracting the background estimation obtained using the proposed method. The difference between Figs. 7(c) and 7(d) is only in that the former uses L_1 minimization at the last step of iterations, while the latter uses L_2 minimization. By comparing the two cases, it can be observed that the L_1 minimization tends to render sparse solutions, where the point sources in the reconstructed distribution show the location and total concentration of the actual objects, while L_2 minimization yields smoother reconstructions. In the case of Fig. 7, the L_2 minimization approach tends to give a more reasonable solution. It should be noted that after subtracting the background signal estimated by the proposed method, we can use any algorithm for object reconstruction. Finally, for comparison Fig. 7(e) shows the result of reconstruction using L_2minimization after subtracting the background estimation obtained based on the uniform background model.

B. Cross Talk between Object and Background Signals and Its Impact on the Performance of the Proposed Method

The method proposed in this paper is developed using the assumption of the object–background model for the fluorophore distribution. Note that no assumption is made regarding the location, size, and number of objects or the level of background heterogeneity; however, the performance of the proposed method is affected by such parameters. Here we explain that, given a pair of subspaces, the main factor affecting the performance of the proposed method in resolving the object and background signals is the level of cross talk between the two.

The extent to which M_o^o and M_b^b approximate, respectively, M_o and M_b determines the reduction gain [defined in Eq. (8)] achieved by the algorithm. A poor representation in the background domain, for instance, limits the amount of the background signal that can be estimated and subtracted. Such poor representation is one of the main reasons that the uniform background model fails to achieve a significant level of background reduction.

On the other hand, the extent to which M_o^b and M_b^o approximate, respectively, M_o and M_b, described here as the cross talk between the object and the background signals, is the main factor affecting the performance of the proposed algorithm in converging to a correct solution. A high level of cross talk may cause the iterative algorithm to converge to a local minimum where the objects are represented partly in the object subspace and partly in the background subspace or vice versa. Furthermore, as shown in Appendix C, a low level of cross talk generally implies faster convergence [the value of the parameter ρ defined in Eq. (C9) indicates the overall level of cross talk between $\bar{\mathbf{M}}_o$ and $\bar{\mathbf{M}}_b$]. The level of cross talk increases as the objects grow in number or size.

The performance of the algorithm in reaching a correct solution is expected to be determined mainly by the level of cross talk rather than by the specific object and background properties, such as their relative locations or shapes; a point further illustrated by a simulation case presented in Fig. 8. In this simulation scenario, we compare the results of object reconstruction for two cases: One for two objects, shown in Fig. 8(a) (for simplicity, only the $z=0$ slice is shown), surrounded by a heterogeneous background, shown in Fig. 8(b); and another one for when the objects are located on the opposite side of the phantom. The object and background relative concentrations are adjusted so that the object and background signals have equal powers. A Gaussian noise with an SNR of 20 dB is added to the measurements. For the former case of objects surrounded by the background, a reduction gain of 17 dB is achieved, and the results of background and object reconstructions are shown in Figs. 8(c) and 8(d), respectively. For the sake of comparison, the result of the reconstruction using the uniform background model is also provided in Fig. 8(e). Figure 8(f) shows the reconstruction result for the latter case of objects located on the opposite side of the phantom (with respect to the circular center) under otherwise the same conditions as the former case. A reduction gain of 15 dB is achieved in this case, which is comparable to the 17 dB reduction gain of the former case. These results confirm the insensitivity of the method to the relative object and background positions. Note that the error in estimation of the background concentration level, as evident by comparing Figs. 8(b) and 8(c), can be justified, as the reconstructed background, Fig. 8(c), has a center of mass that is closer to the surface than the actual background, Fig. 8(b).

This cross talk can be observed in the reconstruction as well, where a large object may partly appear in the background reconstruction. This point is further illustrated in Fig. 9, where a large and diffusive object, Fig. 9(a), surrounded by a highly heterogeneous background, Fig. 9(b), is reconstructed as a point source in the object domain, seen in Fig. 9(c), while making a significant contribution to the reconstructed background, as observed in Fig. 9(d). Interestingly, in this case the background has also con-

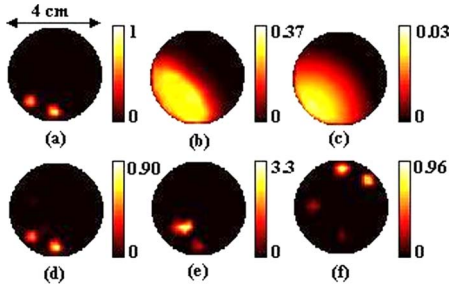

Fig. 8. (Color online) Reconstructions for different object positions relative to the background. (a) Actual object distribution, (b) actual background surrounding the objects, (c) estimated background distribution using the proposed method, (d) estimated object distribution using the proposed method, and (e) estimated object distribution using the uniform background model. The reconstruction result for when objects are located on the opposite side of the phantom, shown in (f), is comparable to (d), which verifies the insensitivity of the proposed method to the location of objects. All slices are at $z=0$ cm.

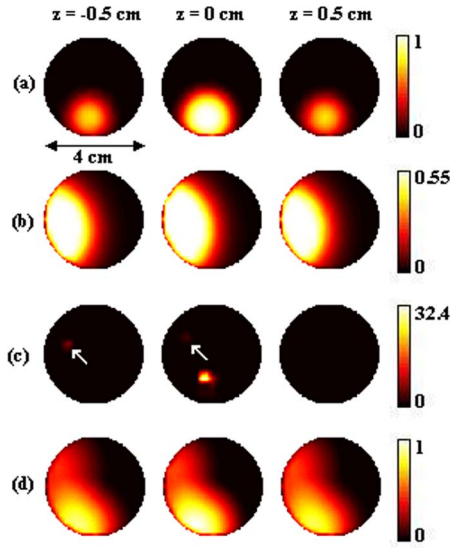

Fig. 9. (Color online) Effect of high object and background signal cross talk observed in the reconstructions. A large object (a) in the presence of the background (b) is reconstructed as a point source in the object domain (c), while significantly contributing to the background estimation (d). Arrows in (c) point to the contribution of the background fluorescence to the object reconstruction.

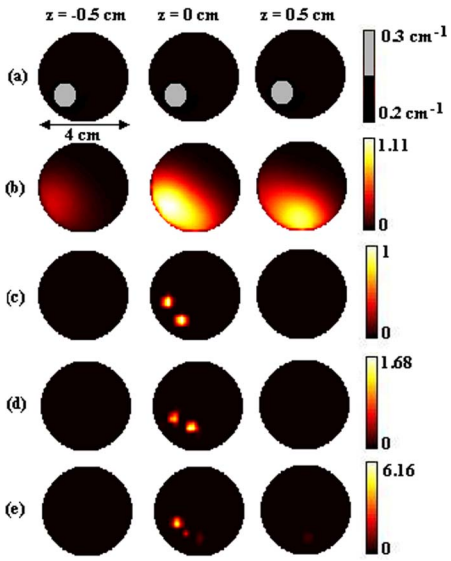

Fig. 10. (Color online) Reconstruction in the presence of optical absorption perturbation depicted in (a). The actual background and object distributions are shown in (b) and (c), respectively, and (d) shows the object distribution estimated using the proposed method. (e) Shows the object estimated using the uniform background model.

tributed to the reconstructed object distribution [pointed to by arrows in Fig. 9(c)]. Finally, note that the similarly large object in Fig. 7 was reconstructed as a collection of point sources using a smoother background subspace than the simulation case in Fig. 9, which implies a lower level of cross talk for the case in Fig. 7. Specifically, the background subspace for the simulation case of Fig. 7 uses $\omega = 1/2\ \mathrm{cm}^{-1}$ [refer to Eq. (6)], while the case of Fig. 9 uses $\omega = 1/3\ \mathrm{cm}^{-1}$. This parameter setting was chosen to demonstrate the cross talk effect in the reconstructions.

C. Reconstruction in the Presence of Optical Perturbations

Quantitative reconstruction of fluorophore concentration is especially complicated by imperfect knowledge about the optical properties of tissue. The actual optical absorption and scattering coefficients of tissue vary spatially and therefore often differ from the assumed values. Such perturbations induce modeling mismatch errors, which affect the performance of reconstruction algorithms. Certain methods, such as Born normalization, can be used to mitigate the impact of modeling mismatch errors. Such approaches are readily applicable to the method proposed in this paper as well.

While our method does not affect the quantitative recovery capability, it allows all existing approaches for such quantification to have better results due to a smaller background fluorescence effect. This fact is further illustrated in Fig. 10. In this simulation case, a cylindrical absorptive perturbation with an absorption coefficient of $0.3\ \mathrm{cm}^{-1}$, shown in Fig. 10(a), is introduced in the phantom, which has an otherwise uniform absorption coefficient of $0.2\ \mathrm{cm}^{-1}$. The proposed method, as in other simulation cases presented in this paper, assumes an optically uniform phantom model and seeks reconstruction of the objects shown in Fig. 10(c) in the presence of the background fluorescence shown in Fig. 10(b). As can be seen in Fig. 10(d), the error resulting from the discrepancies between the assumed and the actual optical properties of tissue does affect the accuracy of the proposed method that is estimating the concentration and location of the objects. However, the proposed method still delivers a more accurate solution, in terms of location and absolute concentration, than the uniform background model, as observed by comparing Figs. 10(d) and 10(e).

7. CONCLUSION

We presented a method for localization of fluorescent objects in the presence of a strong heterogeneous background. The proposed approach partitions the fluorophore distribution into an object and a background distribution, each represented by a different spatial model and expanded using a given set of basis functions. These basis functions are selected to preserve the smooth variations of the background distribution and the sparse distribution of objects. A reconstruction algorithm is used to iteratively solve for the object and background distributions. While conventional methods based on a homogeneous model for the background tend to fail to account for highly heterogeneous distributions, the proposed method produces a more accurate and reliable solution, as observed in various simulation cases presented in this paper.

APPENDIX A: SPATIAL SMOOTHNESS OF BACKGROUND DISTRIBUTIONS AND THE BACKGROUND SUBSPACE

In this section, we present a parameterized definition for spatial smoothness of fluorophore distributions. We also discuss the construction of desirably smooth background subspaces as spanned by smooth basis functions.

The fluorophore distribution over the tissue can be modeled by a nonnegative K-dimensional vector, representing the fluorophore concentration at the kth mesh node. Consider the linear operator $H:\Re^K \rightarrow \Re^K$ acting as a high-pass spatial filter on the fluorophore distribution vectors. The details of the definition of such a high-pass filter are presented in Appendix D, where the filtering operator is represented by a $K \times K$ matrix $\mathbf{H}$. For a given high-pass filter, represented by the matrix $\mathbf{H}$ and a constant $\epsilon \in [0,1]$, we define a given distribution as $(\mathbf{H},\epsilon)$ smooth if it belongs to the set $\bar{\mathbf{S}}_H^{\epsilon}$ defined as

$$\bar{\mathbf{S}}_H^{\epsilon} = \{x \in \Re^{+K} | \ \|\mathbf{H}x\| \leq \epsilon \|x\|\}. \tag{A1}$$

The set $\bar{\mathbf{S}}_H^{\epsilon}$ consists of all distributions whose high-frequency energy content (as determined by the filter $\mathbf{H}$) is not greater than a given fraction, i.e., ϵ, of their total energy. Next we discuss the smoothness of the elements of the set spanned by a set of $P(\mathbf{H},\epsilon)$ smooth basis functions.

Lemma 1. If $\psi_1,\ldots,\psi_P \in \bar{\mathbf{S}}_H^{\epsilon}$ and $\{\varphi_i\}$ is the set of distributions spanned by $(\psi_1,\ldots,\psi_P)$, given as

$$\{\varphi_i\} = \left\{ \sum_{i=1}^{P} a_i \psi_i \,\middle|\, a_i > 0 \right\}, \tag{A2}$$

then $\{\varphi_i\} \subset \bar{\mathbf{S}}_H^{P\epsilon}$.

Proof. Note that if not otherwise stated, all the summation indices in this proof range from 1 to P.

We first show that for $x_1,\ldots,x_P \in \Re^{+K}$, we have

$$\sum_i \|x_i\| \leq P \left\| \sum_i x_i \right\|. \tag{A3}$$

For any $j \in (1,\ldots,P)$, we have

$$\begin{aligned} \left\| \sum_i x_i \right\| &= \sqrt{\left(\sum_i x_i^T\right)\left(\sum_i x_i\right)} \\ &= \sqrt{x_j^T x_j + \underbrace{\sum_{\substack{i,k=1\ldots n \\ (i,k)\neq(j,j)}} x_i^T x_k}_{\geq 0}} \\ &\geqslant \sqrt{x_j^T x_j} = \|x_j\|. \end{aligned} \tag{A4}$$

Summation of the inequality (A4) over $j=1,\ldots,P$ establishes the inequality (A3).

Now for an arbitrary $f \in \{\varphi_i\}$ according to the triangular inequality for the L_2 norm, we have

$$\|\mathbf{H}f\| = \left\| \mathbf{H} \sum_i a_i \psi_i \right\| = \left\| \sum_i a_i \mathbf{H} \psi_i \right\| \leq \sum_i a_i \|\mathbf{H}\psi_i\|. \tag{A5}$$

But since $\psi_1,\ldots,\psi_P \in \bar{\mathbf{S}}_H^{\epsilon}$, we also have

$$\sum_i a_i \|\mathbf{H}\psi_i\| \leq \epsilon \sum_i a_i \|\psi_i\|. \tag{A6}$$

By combining Eqs. (A5) and (A6) and using (A3), we obtain

$$\|\mathbf{H}f\| \leq \epsilon \sum_i a_i \|\psi_i\| \leq \epsilon P \left\| \sum_i a_i \psi_i \right\| = \epsilon P \|f\|, \tag{A7}$$

which shows that for any $f \in \{\varphi_i\}$ we have $f \in \bar{\mathbf{S}}_H^{P\epsilon}$. $\triangle$

The above result demonstrates that if we choose a set of $(\mathbf{H},\epsilon)$ smooth basis functions, then a distribution spanned by these basis functions (using nonnegative expansion coefficients) is $(\mathbf{H},P\epsilon)$ smooth. Since $P>1$, the spanned set is not as smooth as the basis functions; however, if the basis functions are chosen such that $\epsilon \leqslant 1/P$, then the spanned set can be sufficiently smooth as well.

In Subsection 3.A we have used P-modified spatial sinc distributions as the basis functions for construction of the background subspace, as defined by Eqs. (4)–(6). Due to the high spatial smoothness of the sinc functions, ϵ can be chosen to be desirably small, e.g., $\epsilon < 10^{-6}$, if the high-pass filter $\mathbf{H}$ is adjusted to have a minimal spectral overlap with the low-frequency main lobe of the modified sinc function (the value of the ϵ for a modified sinc function is determined based on the sinc parameter ω). It is observed numerically that a few hundred basis functions are sufficient for representation of the background up to a high accuracy. We can adjust $P\epsilon$ to be small enough and achieve desirably smooth background subspaces.

Finally, it is useful to observe that the set $\bar{\mathbf{S}}_H^{\epsilon}$ is, in general, a nonconvex subset of $\Re^K$ (a set $\bar{\mathbf{A}} \subset \Re^K$ is by definition convex if and only if $\forall a \in [0,1]$; $x,y \in \bar{\mathbf{A}} \Rightarrow ax + (1-a)y \in \bar{\mathbf{A}}$). On the other hand, the object subspace, as a set spanned by a linear combination of a set of basis vectors as discussed in Subsection 3.A, is convex. Searching over a convex set results in a convex optimization problem, which is well behaved [30].

APPENDIX B: ORTHOGONAL MATCHING PURSUIT

Consider the problem of estimating a given vector $M \in \Re^{+K}$ as a linear combination of N_o vectors chosen from a set of N vectors $V_i \in \Re^{+K} (i=1,\ldots,N)$. The problem can be formulated as

$$((a_1,\ldots,a_{N_o}), I=(i_1,\ldots,i_{N_o})) = \underset{a_j \geqslant 0, i_j \in \{1,\ldots,N\}}{\arg\min} \left\| \sum_{j=1}^{N_o} a_j V_{i_j} - M \right\|, \tag{B1}$$

where I and $a_j\,(j=1,\ldots,N_o)$ represent, respectively, the set of indices of the chosen vectors and the corresponding expansion coefficients. The optimal solution to this problem requires an exhaustive search, which is computationally prohibitive. The method of OMP [27] seeks a suboptimal solution to the above problem using the iterative search method presented in Table 2 (the method presented in Table 2 is slightly different from the OMP as presented in [27]). In step (a) the residual vector $\tilde{M}$ is set equal to the measurement vector M. In each iteration of the algorithm, the vector that best represents the residual vector $\tilde{M}$ is found in step (b) as the vector with the maximum correlation with the residual vector. In step (c) the set of indices I is augmented with the latest chosen index i_n. In step (d) the set of coefficients that optimally expand

Table 2. Orthogonal Matching Pursuit

(a) Initialization; set $n=1$ and $\tilde{M}=M$.

(b) Choose the next best index $i_n=\arg\max_{i\neq i_1,\ldots,i_{n-1}}\langle\tilde{M},V_i\rangle/(\|\tilde{M}\|\|V_i\|)$.

(c) Form the set of indices $I=(i_1,\ldots,i_n)$.

(d) Estimate the optimal coefficients given the indices $(i_1,\ldots,i_n)$;

$$(a_1,\ldots,a_n)=\arg\min_{a_1,\ldots,a_n\geq 0}\left\|\sum_{j=1}^{n}a_jV_{i_j}-M\right\|.$$

(e) Calculate the residual signal

$$\tilde{M}=M-\sum_{j=1}^{n}a_jV_{i_j}.$$

(f) If $n<N_o$,then $n=n+1$ and go to (b).

the vector M over the set of chosen vectors $(V_{i_1},\ldots,V_{i_n})$ are found, and the new residual vector is calculated in step (e). The iteration repeats until the set I has N_o elements.

The optimization in step (d) usually involves a very low number of scalar variables (up to N_o, at most) and is computationally negligible due to the relatively low value of N_o. Step (b) is the most computationally expensive step, where the next best index is found based on the correlation with all candidate vectors. Compared to the optimizations involved in the other steps of the method proposed in the paper, OMP has negligible complexity.

APPENDIX C: CONVERGENCE OF THE PROPOSED METHOD

In Section 5, a solution for separating the object and background signals in the measurements was presented. This problem can be regarded, in general, as one of describing a vector as a summation of two vectors from two respective sets. This section investigates this problem in its general form and discusses the convergence of the proposed iterative solution.

Specifically, let $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ be two subsets of $\Re^{+N}$ with the following properties:

$$\forall\alpha\geq 0,\quad x\in\bar{\mathbf{A}}\Rightarrow\alpha x\in\bar{\mathbf{A}},\tag{C1}$$

$$\forall\alpha\geq 0,\quad x\in\bar{\mathbf{B}}\Rightarrow\alpha x\in\bar{\mathbf{B}}.\tag{C2}$$

Also, let m be a vector in $\Re^{+N}$. Consider the problem of optimally describing m as a summation of a member of $\bar{\mathbf{A}}$ and a member of $\bar{\mathbf{B}}$:

$$(a_o,b_o)=\arg\min_{a\in\bar{\mathbf{A}},b\in\bar{\mathbf{B}}}\|m-a-b\|.\tag{C3}$$

This paper deals with a special case of Eq. (C3), in which m, a_o, and b_o represent, respectively, the measurement vector M, the object signal M_o, and the background signal M_b, defined as $M_b=\mathbf{Z}X_b$ and $M_o=\mathbf{Z}X_o$. Next we present a framework for representation of the object and background distributions. The subsets $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ correspond to, respectively, the object and background subspaces mapped into $\Re^{+K}$ using the system matrix $\mathbf{Z}$.

Table 3. Iterative Method for Joint Estimation over Two Arbitrary Subspaces

(a) Initialization; set $i=0$, $a_0=\mathbf{0}_N$, and $b_0=\mathbf{0}_N$.

(b) Find the best vector in $\bar{\mathbf{A}}$ given b_i;

$$a_{i+1}=\arg\min_{x\in\bar{\mathbf{A}}}\|m-b_i-x\|.$$

(c) Find the best vector in $\bar{\mathbf{B}}$ given a_{i+1};

$$b_{i+1}=\arg\min_{x\in\bar{\mathbf{B}}}\|m-a_{i+1}-x\|.$$

(d) Calculate the estimation error $E_{i+1}=\|m-a_{i+1}-b_{i+1}\|$.

(e) If the estimation error E_i has converged,[a] then $i=i+1$ and go to (b).

(f) Set $a_o=a_{i+1}$ and $b_o=b_{i+1}$.

[a]Since the estimation error E_i is positive and decreasing, it is converging (see Appendix C).

The iterative method presented in Table 3 seeks a suboptimal solution to Eq. (C3) by sequentially refining the estimation of each of the vectors a and b using the current estimation of the other, as performed in steps (b) and (c). The estimation error E_i is used to test the convergence between iterations. In this appendix, it is shown that the iterative algorithm presented in the general form in Table 3 converges according to the specified convergence criteria.

Due to definition of a_{i+1} given in step (b), we have

$$\forall x\in\bar{\mathbf{A}};\|m-b_i-a_{i+1}\|\leq\|m-b_i-x\|.\tag{C4}$$

Therefore

$$\|m-b_i-a_{i+1}\|\leq\|m-b_i-a_i\|.\tag{C5}$$

Similarly, using the definition of b_{i+1} given in step (c), we have

$$\|m-a_{i+1}-b_{i+1}\|\leq\|m-a_{i+1}-b_i\|.\tag{C6}$$

By combining the inequalities (C5) and (C6), we get

$$\|m-a_{i+1}-b_{i+1}\|\leq\|m-a_i-b_i\|,\tag{C7}$$

or

$$E_{i+1}\leq E_i.\tag{C8}$$

Since, according to inequality (C8) E_i is decreasing versus i and positive, it also converges. While the above argument proves that the algorithm presented in Table 3 converges, as inspected in step (e), it does not establish the convergence or stability of a_i or b_i.

The iterative method proposed in Table 1 can be viewed in the OMP mode as a special form of the iterative algorithm above. If OMP provided an optimal solution, then the results suggested by inequality (C8) would establish the convergence in the OMP mode. Although OMP is a suboptimal solver, the simulation results presented in Section 5 verify the applicability of the above results. Since the actual optimizations used in the iterative algorithm presented in Table 1 do not always converge to the optimal solutions, the estimation error, while following a decreasing trend, might not necessarily be decreasing; accordingly, our algorithm tests the level of fluctuations of

the error in the last few iterations. Next we discuss the convergence rate of the algorithm presented in Table 3. Define the distance between the two subsets $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ as

$$\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}} = \sup_{a\in\bar{\mathbf{A}},b\in\bar{\mathbf{B}}} \frac{|\langle a,b\rangle|}{\|a\|\cdot\|b\|}. \tag{C9}$$

For a vector x, define $\hat{x}\equiv x/\|x\|$. Due to the definitions of a_{i+1} and b_i and also to the closeness of the subsets $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ over scalar multiplication [as given in Eqs. (C1) and (C2)], we have

$$\begin{aligned}\|a_{i+1}\| &= |\langle m-b_i,\hat{a}_{i+1}\rangle| \\ &= |\langle m-\langle m-a_i,\hat{b}_i\rangle\hat{b}_i,\hat{a}_{i+1}\rangle| \\ &= \left|\underbrace{\langle m,\hat{a}_{i+1}\rangle-\langle m,\hat{b}_i\rangle\langle\hat{b}_i,\hat{a}_{i+1}\rangle}_{\equiv c_i}+\langle a_i,\hat{b}_i\rangle\underbrace{\langle\hat{b}_i,\hat{a}_{i+1}\rangle}_{\equiv\lambda_i}\right|,\end{aligned} \tag{C10}$$

or, in other terms, given the definitions of c_i and λ_i in Eq. (C10) above, we can write

$$\|a_{i+1}\| = |c_i+\lambda_i\langle a_i,\hat{b}_i\rangle|. \tag{C11}$$

This equation suggests that the upper limit of λ_i, i.e., $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}$, determines the rate of the convergence. Equation (C11) suggests the value $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}$ affects the rate of the algorithm convergence, as a decreasing $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}$ hastens the convergence. Erratic behavior is expected for $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}=1$. Analytical solutions for the above iterative approach can be obtained for very simple cases, such as when each of the subsets $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ is spanned using only one unit vector. It can be seen that for such a case, the rate of the convergence is explicitly determined by the normalized inner product of the unit vector of $\bar{\mathbf{A}}$ and the unit vector of $\bar{\mathbf{B}}$.

For the OMP mode, the object subspace consists of distribution vectors with at most N_o nonzero elements. The distance between the object subspace and the background subspace is relatively small. For example, for the set of parameters used in the case study presented in Section 5, namely, $N_o=2$, $\omega=1/3\ \mathrm{cm}^{-1}$, and $P=220$, we have $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}=0.6$. As suggested by Eq. (C11), due to the small value of $\rho_{\bar{\mathbf{A}}\bar{\mathbf{B}}}$ (as compared to 1), the algorithm is expected to converge. On the other hand, for the BP mode, where the sparseness of the object term is enforced by minimization of its L_1 norm during reconstruction, the object subspace is essentially the entire space $\Re^{+K}$, as there is no upper limit on the number of nonzero elements. In this sense, the distance between the object and the background subspaces is one; therefore, the estimation error E_i will converge for the OMP mode, although it might show erratic behavior in the BP mode. This erratic behavior justifies not initially enforcing sparseness through L_1 minimization. Simulation results suggest that L_1 minimization can be reliably used to improve the estimations once a reliable estimation of the background is found in the OMP mode.

APPENDIX D: SPATIAL FILTERING OF DISTRIBUTION VECTORS

In this appendix, the spatial filtering of a given K-dimensional distribution vector (where K is the number of nodes in the tetrahedral FEM mesh), used in Appendix A to establish a criterion for spatial smoothness, is described. Because of the irregularity of the locations of the mesh nodes in the three-dimensional space, spatial filtering, though conceptually simple, needs to be clarified. Here we explain how to construct a filtering operator $H:\Re^K\rightarrow\Re^K$ from a given three-dimensional digital filter.

Suppose $x\in\Re^{+K}$ is an arbitrary distribution over the K mesh nodes, where each element of x represents the concentration of fluorophores on the corresponding mesh node. Consider an $N_x\times N_y\times N_z$ rectangular grid in $\Re^3$ encompassing the entire tetrahedral mesh. The distribution x, defined over the mesh, can be mapped into a distribution over the rectangular grid, denoted by an $(N_xN_yN_z)$-dimensional vector y, using an interpolation operator $I:\Re^{+K}\rightarrow\Re^{+N_xN_yN_z}$, also representable by an $N_xN_yN_z\times K$ matrix $\mathbf{I}$ such that $y=\mathbf{I}x$. One simple interpolation scheme describes the concentration on each grid point as a linear combination of the concentrations on the four vertices of the tetrahedral element that contains that grid point. If the rectangular grid is chosen such that $N_xN_yN_z>K$, then the inverse of the interpolation operator is given by a $K\times N_xN_yN_z$ matrix $\mathbf{J}$ (defined as $\mathbf{J}=(\mathbf{I}^T\mathbf{I})^{-1}\mathbf{I}^T$).

Consider a spatial discrete-time filter, represented by the three-dimensional sequence $g(m,n,p)$, where m,n,p are integers. This filter can be represented by an $N_xN_yN_z\times N_xN_yN_z$ matrix $\mathbf{G}$ [31] operating on the grid distribution y such that the filtered grid distribution is given by $\mathbf{G}y$. The filtered mesh distribution is defined here as the filtered grid distribution mapped back to the mesh by the inverse interpolation operator $\mathbf{J}$. In other words, the filtering matrix for the tetrahedral mesh $\mathbf{H}$ is given as $\mathbf{H}=(\mathbf{I}^T\mathbf{I})^{-1}\mathbf{I}^T\mathbf{G}\mathbf{I}$, which is the grid filter matrix $\mathbf{G}$ interpolated over the tetrahedral mesh.

ACKNOWLEDGMENTS

This work was supported by the David and Lucile Packard Foundation. The authors thank Wael Yared, Nara Narayanan, and Jeffrey D. Peterson at VisEn Medical Inc., Woburn Massachusetts, and Charles H. Camp, Jr. at Georgia Institute of Technology, Atlanta, Georgia, for their helpful discussions.

REFERENCES

1. K. Licha, "Contrast agents for optical imaging," Top. Curr. Chem. **222**, 1–29 (2002).
2. R. Weissleder, "A clearer vision for *in vivo* imaging," Nat. Biotechnol. **19**, 316–317 (2001).
3. V. Ntziachristos, C. H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity *in vivo*," Nat. Med. (N.Y.) **8**, 757–761 (2002).
4. E. Chang, J. Sun, J. S. Miller, W. W. Yu, V. L. Colvin, J. L. West, and R. Drezek, "Protease-activated quantum dot probes," Proc. SPIE **6191**, 61911E.1–10 (2006).
5. C. Bremer, C. H. Tung, and R. Weissleder, "*In vivo* molecular target assessment of matrix metalloproteinase inhibition," Nat. Med. (N.Y.) **7**, 743–748 (2001).
6. R. Weissleder, "Molecular imaging in cancer," Science **312**, 1168–1171 (2006).

7. G. Zacharakis, J. Ripoll, R. Weissleder, and V. Ntziachristos, "Fluorescent protein tomography scanner for small animal imaging," IEEE Trans. Med. Imaging **24**, 878–885 (2005).
8. E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution fluorescence molecular imaging system for small animal imaging," Med. Phys. **30**, 901–911 (2003).
9. A. Godavarty, M. J. Eppstein, Z. Chaoyang, S. Theru, A. B. Thompson, M. Gurfinkel, and E. M. Sevick-Muraca, "Fluorescence-enhanced optical imaging in large tissue volumes using a gain-modulated ICCD camera," Phys. Med. Biol. **48**, 1701–1720 (2003).
10. H. Jiang, "Frequency-domain fluorescent diffusion tomography: A finite-element-based algorithm and simulations," Appl. Opt. **37**, 5337–5343 (1998).
11. A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three-dimensional *in vivo* fluorescence diffuse optical tomography of breast cancer in humans," Opt. Express **15**, 6696–6716 (2007).
12. A. Soubret and V. Ntziachristos, "Fluorescence molecular tomography in the presence of background fluorescence," Phys. Med. Biol. **51**, 3983–4001 (2006).
13. A. K. Sahu, R. Roy, A. Joshi, and E. M. Sevick-Muraca, "Evaluation of anatomical structure and non-uniform distribution of imaging agent in near-infrared fluorescence-enhanced optical tomography," Opt. Express **13**, 10182–10199 (2005).
14. J. Chang, H. L. Graber, and R. L. Barbour, "Improved reconstruction algorithm for luminescence optical tomography when background lumiphore is present," Appl. Opt. **37**, 3547–3552 (1998).
15. P. Mohajerani, A. A. Eftekhar, J. Huang, and A. Adibi, "Optimal sparse solution for fluorescent diffuse optical tomography: Theory and phantom experimental results," Appl. Opt. **46**, 1679–1685 (2007).
16. A. Godavarty, M. J. Eppstein, C. Zhang, and E. M. Sevick-Muraca, "Detection of single and multiple targets in tissue phantoms with fluorescence-enhanced optical imaging: Feasibility study," Radiology **235**, 148–154 (2005).
17. E. L. Hull, M. G. Nichols, and T. H. Foster, "Localization of luminescent inhomogeneities in turbid media with spatially resolved measurements of cw diffuse luminescence emittance," Appl. Opt. **37**, 2755–2765 (1998).
18. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," Phys. Med. Biol. **50**, 1–43 (2005).
19. A. Cong and G. Wang, "A finite-element-based reconstruction method for 3D fluorescence tomography," Opt. Express **13**, 9847–9857 (2005).
20. S. R. Arridge, M. Schweiger, M. Hiraoka, and D. T. Delpy, "A finite element approach for modeling photon transport in tissue," Med. Phys. **20**, 299–309 (1993).
21. A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," Appl. Opt. **42**, 3081–3094 (2003).
22. R. B. Schulz, J. Ripoll, and V. Ntziachristos, "Experimental fluorescence tomography of tissues with noncontact measurements," IEEE Trans. Med. Imaging **23**, 492–500 (2004).
23. A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized Bornratio," IEEE Trans. Med. Imaging **24**, 1377–1386 (2005).
24. R. Roy, A. Godavarty, and E. M. Sevick-Muraca, "Fluorescence-enhanced optical tomography of a large tissue phantom using point illumination geometries," J. Biomed. Opt. **11**, 044007 (2006).
25. A. B. Milstein, "Imaging of near-infrared fluorescence, absorption, and scattering in turbid media," Ph.D. thesis, (Purdue University, 2004).
26. S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **20**, 33–61 (1998).
27. Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proceedings of the 27th Asilomar Conference on Signals Systems and Computers* (IEEE, 1993), pp. 40–44.
28. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," Opt. Lett. **26**, 893–895 (2001).
29. M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy, "The finite element method for the propagation of light in scattering media: Boundary and source conditions," Med. Phys. **22**, 1779–1792 (1995).
30. S. P. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge U. Press, 2004).
31. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, 1989).